Bioimaging 4 (1996) 73–83. Printed in the UK

High sensitivity, high resolution physical mapping by fluorescence *in situ* hybridization on to individual straightened DNA molecules

Mei Wang†, Thomas Duell†‡, Joe W Gray†§ and Heinz-Ulrich G Weier†‖

† Resource for Molecular Cytogenetics, Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA
‡ Medizinische Klinik III, Ludwig-Maximilians-University, D-81377 Muenchen, Germany
§ Division of Molecular Cytometry, University of California, San Francisco, CA 94143, USA

Submitted 11 June 1996, accepted 2 October 1996

Abstract. High resolution physical mapping of clonal DNA fragments with kilobase (kb) resolution can now be performed rapidly by fluorescence *in situ* hybridization (FISH) onto individual DNA molecules (DNA fibers). We developed a sensitive procedure termed 'quantitative DNA fiber mapping' which consists of three steps: preparation of DNA fibers, hybridization of non-isotopically labeled probes and determination of the relative mapping position by fluorescence image analysis. The DNA fibers are produced by binding linearized DNA molecules with one or both ends to a solid substrate followed by homogeneous stretching of the molecule by the action of a receding meniscus during drying ('molecular combing'). In a slight variation of this protocol, we deposit circular DNA molecules. Substrates for DNA immobilization are glass slides, coverslips or thin sheets of mica derivatized with amino-silane. Probes are prepared to counterstain the DNA fibers, from the clones to be mapped and for specific landmarks along linear or circular DNA molecules such as cloning vector sequences. Following hybridization and immunocytochemical detection of bound probes, images are analysed and relative distances are recorded for map assembly. Here, we describe our experience with substrate preparation, molecular combing and mapping of cloned or enzymatically synthesized probes ranging in size from 1.2 kb to 100 kb along DNA molecules that are between 17 kb and 1200 kb in size.

Keywords: DNA molecules, molecular combing, physical mapping, fluorescence *in situ* hybridization (FISH), image analysis

1. Introduction

Elegant yet simple, FISH has proven its utility for enumeration and localization of specific DNA sequences in individual somatic cell interphase nuclei, tissue sections and germ cells many times. The chromatin organization in interphase cell nuclei, however, severely limits the resolution with which two linked probes can be mapped relative to one another. Chromatin organization also impacts on sensitivity and small targets may not be detected when the DNA target accessibility for the probes is poor.

Some of these problems can at least partially be overcome by releasing the DNA from cell nuclei by mechanical or chemical means (Heng *et al* 1992, Lawrence *et al* 1992, Wiegant *et al* 1992, Parra and Windle 1993, Haaf and Ward 1994). Techniques such as the recently described DIRVISH or DNA halo preparations release the chromatin from interphase cells and deposit the decondensed genomic DNA in a more or less random fashion on glass slides or other suitable substrates (Parra

‖ To whom correspondence should be addressed. E-mail address: weier@white.lbl.gov

0966-9051/96/020073+11$19.50

and Windle 1993, Tocharoentanaphol *et al* 1994, Suto *et al* 1996). The DNA, now much more accessible to the probes, appears in bundles or fibers of DNA which are extended to a certain, often highly variable, degree. Such DNA fibers have been used successfully to demonstrate overlap between clones or genomic deletions (Florijn *et al* 1995, Suto *et al* 1996). The term 'fiber' had originally been used to describe the 30 nm fibers of chromatin. The publications mentioned above do not differentiate well between fibers of chromatin or pure DNA and refer to either one as 'DNA fibers'. Since this has become a popular term in a relatively short time, we will use it in this paper to describe individual DNA molecules.

One of the major steps in position cloning of disease genes is the construction of ordered high resolution physical maps. This requires knowledge of the extent and orientation of clones in such a map. The application of the before mentioned techniques for assembly of such high resolution physical maps, however, is rather time consuming, because the genomic templates contain mostly non-target DNA sequences and hybridization targets suitable for analysis are rare and, thus, hard to find in the useful area of a microscopic slide. The existence of duplicated genomic regions, gene families or pseudogenes which map to different autosomes further complicates the analysis of hybridization pattern, because these will produce numerous disperse hybridization domains rather than a single unique hybridization signal.

More recently, techniques for mapping of cloned probes on DNA fragments enriched by pulsed field gel electrophoresis (PFGE) have been described that provide a much higher density of target DNA molecules and, at the same time, limit the complexity of the hybridization target so that some of the before mentioned cross-hybridization problems may be circumvented (Heiskanen *et al* 1994, Rosenberg *et al* 1995). Mechanical deposition of DNA molecules however, does not allow precise control of the extent of DNA stretching so that standards are needed for normalization and distance measurements (Heiskanen *et al* 1994).

It was recently demonstrated that the hydrodynamic action of a receding meniscus during drying (termed 'molecular combing') resulted in homogeneously stretched DNA of approximately 2.3 kb μm^{-1}, i.e. extended molecules which are about 30% longer than relaxed DNA of the same size (Bensimon *et al* 1994, Weier *et al* 1995b). We observed reproducible stretching with DNA molecules ranging in size from 48 kb to 490 kb (Weier *et al* 1995b), which were bound to a solid substrate at either one or both of their respective ends. These homogeneously stretched DNA molecules were then used as targets for probe mapping using FISH. To reflect its quantitative character, we termed the procedure 'quantitative DNA fiber mapping (QDFM)' (Weier *et al* 1995b). Applying this procedure to the measurement of overlap between cosmid clones and collinear P1 DNA molecules required separation of the P1 molecules from the bacterial host, linearization of the P1 DNA by digestion with a rare cutting enzyme and PFGE purification of the linear P1 DNA molecules. The linearized and purified P1 molecules were then bound to silanated glass slides, subjected to molecular combing and used as mapping templates (Weier *et al* 1995b). This allowed us to measure the extent as well as the orientation of overlap between clones, because the vector sequences remained attached to the DNA fiber and provided additional reference points.

Linearization of circular DNA molecules from clones from large insert DNA libraries such as the P1 (Shepherd *et al* 1994), PAC (Ioannou *et al* 1994) or BAC (Shizuya *et al* 1992) libraries is not without complications. The P1 and PAC systems provide a unique Not 1 site in the cloning vector. However, linearization with this enzyme is complicated, when the insert contains a Not 1 site. Furthermore, some cloning vectors (such as the cosmid vector pWE15 and the BAC vector) were engineered to facilitate purification of the cloned insert and contain Not 1 sites on either side of the cloning site. When using such clones for QDFM, digestion with Not 1 is not useful, because it would separate vector and insert sequences. We therefore decided to investigate deposition of circular, uncut BAC or P1 DNA or randomly broken molecules on to the solid substrate and performed molecular combing and physical mapping on to these molecules as described for enzymatically linearized DNA molecules.

In the following, we summarize our experience with preparation of glass substrates and high molecular weight DNA and QDFM. The examples were selected to represent a broad range of applications ranging from map assembly to hybridization on to linear and circular DNA molecules in support of large scale sequencing efforts.

2. Materials and methods

2.1. Preparation of high molecular weight DNA

High molecular weight P1 DNA as well as plasmid, cosmid and BAC DNA is isolated routinely in our laboratory using an alkaline lysis protocol (Birnboim and Doly 1979, Weier *et al* 1995a). Prior to sizing by pulsed field electrophoresis, high molecular weight DNA molecules were linearized by digestion of approximately 1 μg of P1 DNA with 10U Not 1 or Sfi I for 1 hour at 37 °C or 50 °C, respectively. The DNA was then loaded on a 1.0% low melting point agarose gel and electrophoresed for 15 h using a BioRad CHEF mapper according to the manufacturer (BioRad, Hercules, CA) (Weier *et al* 1995a). The band containing the linear P1 DNA (figure 1A) was then excised from the gel, the slice was equilibrated with agarase buffer and the gel was digested with agarase as recommended by the supplier (New England Biolabs, Beverly, MA).

High molecular DNA fibers from BAC clones were prepared from circular and randomly broken DNA molecules. Following alkaline lysis, phenol/chloroform

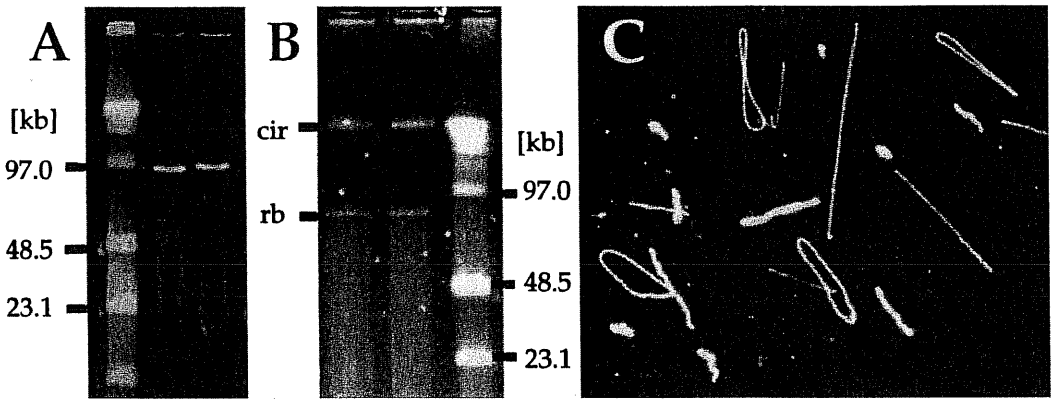

Figure 1. Pulsed field gel electrophoretic (PFGE) and microscopic analysis of P1 DNA molecules (mapping templates). (A) Enzymatic restriction with Not 1 cuts most P1 clones at a unique site within the vector. Linear molecules appear as a single band on the PFGE gel. The sizes of bands in the sizemarker lane (left) are indicated. The two lanes to the right each contain 1 μg of Not 1 digested DNA from two different P1 clones. (B) The alkaline lysis procedure generates circular and randomly broken or sheared DNA molecules. This gel photograph shows the migration pattern observed with undigested P1 DNA molecules. Circular molecules (cir) migrate much slower than full length linear molecules randomly broken during the preparation (rb). Both lanes contain 1 μg of the same P1 DNA clone. The sizemarker lanes in A and B contain a lambda DNA concatemer. (C) Circular DNA molecules are isolated by agarose digestion of PFGE gel slices containing the desired band. The individual molecules in these preparations can be visualized in the fluorescence microscope after staining with YOYO. The preparations typically contain circular molecules and varying amounts of randomly broken DNA molecules.

extraction and ethanol precipitation (Weier *et al* 1995b), the DNA was loaded on the PFGE gel without digestion. Bands corresponding to closed circular and randomly broken full length molecules (figure 1B) were excised and digested with agarase.

High molecular weight YAC DNA from clones '141G6' and '35B2' was isolated by PFGE as described before (Weier *et al* 1995b). The human DNA inserts in YAC clones 141G6 and 35B2 have a size of $\sim$ 475 kb and $\sim$ 290 kb, respectively (Patil *et al* 1994). We excised gel slices containing the YACs from the low melting point agarose, digested the agarose with agarase and equilibrated YAC DNA with 100 mM NaCl (Weier *et al* 1995b). All high molecular weight DNAs were stored at 4 °C.

The integrity of the prepared DNA molecules was estimated by staining small aliquots of DNA (typically 1 μl) with 0.5 μM YOYO-1 (Molecular Probes, Eugene, OR) (Glazer and Rye 1992) and resuspending them in 1% *p*-phenylenediamine, 15 mM NaCl, 1 mM H_2PO_4, pH 8.0, 90% glycerol (AF-solution) (Johnson and Nogueira Araujo 1981) or double distilled water. In some cases, this suspension was further diluted with water to provide a suitable DNA concentration, before it was placed on a microscope slide and covered with a coverslip. If the DNA concentration was low, for example, when a relatively large gel slice had to be digested, DNA stained with YOYO was placed undiluted on the slide. The DNA molecules were then observed on a Zeiss Axioskope microscope using a filter set for FITC (Chroma Technology, Brattleboro, VT) and a standard HBO 100 light source (figure 1C).

2.2. Substrate preparation and molecular combing

Glass slides were cleaned mechanically and rinsed with double distilled water to remove any residual glass particles. After boiling in distilled water for 10 minutes and air drying, the slides were immersed in 18 M sulfuric acid (TJ Baker, Phillipsburg, NJ) for 30–40 minutes to remove any organic residues. After another boiling step in double distilled water, the slides were air dried and stored under dust free conditions until further use.

Silane modification was performed using 3-aminopropyltriethoxy-silane (APS, Sigma, St. Louis, MO). Pretreated dry slides were immersed for 10 minutes in a solution of 0.1% APS in 95% ethanol. Following removal of slides from the silane solution, they were intensively rinsed with double distilled water, immersed in double distilled water for 2 minutes and dehydrated by immersing in absolute ethanol. After drying at 65 °C on a heat plate in upright position, they were baked for another 10 minutes at this temperature. Slides were stored at 4 °C in a sealed slide box under a nitrogen atmosphere and aged for 2–6 weeks to provide optimal results.

Prior to molecular combing, the purified P1, BAC or YAC DNA molecules were diluted with water to about 10 pg μl^{-1} (equivalent to several thousand molecules per μl) and mixed with an equal amount of YOYO (10^{-5} M). Two microliters of this DNA were placed on a non-treated 22 mm $\times$ 22 mm coverslip. An APS-treated slide was then placed on top and the liquid was allowed to dry at 4 °C or 20 °C overnight. In some experiments, we added 8 μl AF-solution to 2 μl of YOYO stained DNA molecules to minimize breakage before further dilution with

water. Since this resulted in further dilution of the DNA molecules, it was applied only to samples with relatively high DNA concentration ($\sim$ 1 ng μl^{-1}).

2.3. Probe preparation and labeling

The P1 clone '1143' was isolated from a human genomic P1 library (Shepherd *et al* 1994) in the course of assembly of a physical map based on STS content mapping (J F Cheng, unpublished). The size of the human genomic insert in this P1 clone ($\sim$ 80 kb) was calculated from the PFGE analysis of Not I-linearized P1 molecules. The size of the recombinant cloning vector pAd10-SacBII is $\sim$ 17 kb. Other locus-specific large insert DNA clones from human P1 or BAC libraries were isolated by PCR screening (Weier *et al* 1995b). Clones from the Lawrence Livermore National Library (LLNL) chromosome 22-specific cosmid library were kindly provided by Dr U-J Kim, California Institute of Technology, Pasadena. Cosmids used in the studies reported here (31F3 and 52F2) map to the Ig lambda variable region gene cluster on the long arm of chromosome 22 (Kawasaki *et al* 1995). All cosmid, P1 or BAC probes were generated from DNA prepared from overnight cultures by an alkaline lysis protocol (Weier *et al* 1995b). The YAC DNA for probe preparation was isolated from whole yeast cultures following standard procedures (Sherman *et al* 1986). Plasmid probes, used to mark the ends of YAC molecules, contain either the centric or acentric YAC arm (pJS97 and pJS98, respectively) or both (pYAC3). The DNA from these clones was also prepared by alkaline lysis.

Aliquots of 400 ng of DNA from the P1, BAC and YAC clones were labeled by random priming incorporating biotin-dUTP or digoxigenin (dig.)-dUTP (Weier *et al* 1995b). The DNAs from plasmid clones were labeled by random priming incorporating digoxigenin-dUTP (Weier *et al* 1995a). The chromosome 22-specific cosmid clones were random primed in the presence of digoxigenin-dUTP (clone 31F3) or biotin-dUTP (clone 52F2). A biotinylated DNA probe for counterstaining molecules of YAC clone '141G6' was made by random priming of DNA isolated from the YAC containing yeast clone. Probe for the P1 cloning vector was prepared from plasmid pAd10-SacBII DNA by random priming in the presence of FITC-dUTP (Weier *et al* 1995b).

Furthermore, we applied *in vitro* DNA amplification using the polymerase chain reaction to prepare probes that bind specifically to the BAC cloning vector (Weier, manuscript in preparation). We designed five sets of oligonucleotide primer pairs that flank 1100–1400 bp of minimally overlapping BAC vector sequences. The PCR products were phenol/chloroform extracted, ethanol precipitated and labeled by random priming. One of the five products was labeled with digoxigenin and the other four products were fluoresceinated, so that the digoxigenin-labeled probe binds at position 4319–5703 of the $\sim$ 7.4 kb BAC vector. We also generated several probes of $\sim$ 1200–1400 bp each for the P1 cloning vector by *in vitro* DNA amplification. These probes can be combined with recombinant (plasmid pAd10-SacBII) P1 vector probe to determine the localization, extent and orientation of the P1 vector part of DNA fibers.

2.4. FISH

When mapping on to YAC DNA fibers, digoxigenin-labeled DNA from pYAC3, pJS97 or pJS98 was included in the hybridization probe mixture to specifically label the ends of untruncated YAC molecules.

We followed two slightly different hybridization schemes employing either two (red, green) or three (blue, red, green) different colors. Biotinylated probes can be detected in green or blue using either avidin-FITC (Vector, Burlingame, CA) or -AMCA (Vector). Digoxigenin-labeled probes were detected with rhodamine-conjugated antibodies (Boehringer Mannheim, Indianapolis, IN)(red), while the FITC-labeled probes were detected with antibodies emitting green fluorescence.

Seven microliters of hybridization 'master mix' comprised of 78.6% formamide, 14.3% dextran sulfate in 2.9×SSC, pH 7.0 (2×SSC is 300 mM NaCl, 30 mM Na citrate) were mixed with 1 μg human Cot-1 DNA (Gibco/LTI) and approximately 20 ng of each probe, applied to the slide and coverslipped. Denaturation was done at 100 °C for 2–3 min on a heat block, and the probe was allowed to hybridize overnight at 37 °C. In some experiments, we pre-denatured the probes for 10 min at 92 °C, applied them to the slides, coverslipped and heated the slide for 40 s to 92 °C. The slides were then washed in three changes of 2×SSC at 20 °C for 15 min each with slight agitation. Bound probes were detected by conjugation with avidin or antibodies and hybridization signals were amplified with biotinylated goat-anti-avidin (Vector) followed by a second layer of avidin and a Texas Red-labeled rabbit antibody against sheep IgG (Vector). In the triple color detection scheme, detection of FITC labeled probes was performed with a mouse anti-FITC antibody (DAKO, Carpintera, CA) followed by incubation with an FITC-conjugated horse-anti-mouse antibody (Vector). Slides were finally washed twice in 2×SSC and mounted in AF-solution for microscopic inspection.

2.5. Image acquisition and analysis

Images were acquired on a quantitative image processing system based on a Zeiss fluorescence microscope equipped with 63×, 1.25 NA and 40×, 1.3 NA oil objectives, a Photometrics cooled CCD camera, multiband pass filters for simultaneous observation of FITC and Texas Red or AMCA/DAPI (Chroma Technology, Brattleboro, VT) and a SUN SPARC workstation (Mascio *et al* 1995).

Images from P1/BAC and YAC molecules were recorded on the SUN system, converted to 24-bit tiff-format images, transferred to an Apple Macintosh computer

and analysed using routines of 'NIH Image V1.57'. This software allows the user to interactively trace DNA fibers by drawing a segmented line and then calculates the length of the line in pixels. The pixel spacing was known from the microscope objective used in the experiment (we use 63× magnification for molecules up to 100 kb, 40× magnification for all larger molecules) and was converted into μm (or kb using a factor of 2.3 kb μm^{-1}) (Weier *et al* 1995b). We measured several distances along the hybridized fibers in triplicate. On images recorded from P1 fibers, we measured the total length of the fiber, the size of its vector part, the size of probe hybridization domains as well as distances from either end.

The output of NIH Image provides the measurements in the form of a list. This list was imported into Microsoft Excel, reformatted manually and used to calculate average values for each fiber. The average values of several fibers were used to calculate mean mapping values and standard deviations for fibers measured in one experiment.

3. Results and discussion

3.1. Preparation of glass substrates and molecular combing

The derivatization of glass surface with silane is one of the most critical steps in our procedure. Different reactive groups allow coupling of the DNA molecules to the solid surface. Bensimon *et al* (1994), for example, report molecular combing of lambda and E. coli DNA using coverslips that carried a reactive vinylgroup on the surface. We found that these coverslips gave high levels of background fluorescence in FISH experiments, and derivatized our solid substrates with a silane that carries a covalently linked aminogroup (APS) (Weier *et al* 1995b). The same silane has been used to prepare derivatized glass coverslips and sheets of mica for molecular combing (Hu *et al* 1996).

The slide preparation has not yet been standardized and appears to depend on the batch of glass used in the manufacturing process. Great differences exist between slides and coverslips, the former typically made from soda lime glass, while the latter are typically prepared from borosilicate. In our hands, coverslips need more silanation to bind DNA to the same effect, and we increase the silane concentration to up to 0.5%. We also noted that glass slides from different manufacturers and even from the same brand may yield DNA fibers which, to varying degrees, lose DNA during the denaturation and hybridization procedure. Consequently, we prepare small batches of ~ 50 slides, and test the DNA binding and QDFM performance on 2–3 slides with DNA molecules of known integrity. A good batch of slides binds most of the DNA molecules on one or both ends within a few minutes of the APS-treated slide being placed on the top of the coverslip carrying the DNA.

Molecular combing done at room temperature (20 °C) or at 4 °C in the refrigerator did not show consistent differences. Drying at 20 °C and ~ 50–55% relative humidity is typically somewhat faster than molecular combing in the refrigerator, but often slides are left to dry overnight. The observed stretching to ~ 2.3 kb μm^{-1} appears to be an upper limit for the DNA fibers and higher force seems to break the molecules. We found, however, areas on the slides that showed an increased number of incompletely stretched DNA molecules. Such incompletely stretched molecules are identified by their strikingly different appearance: they do not display a bead-like structure after FISH and exhibit strong fluorescence signals while showing significantly shorter signal domains. These molecules are easily identified using probes that bind to 'constant' regions such as the vector part of a recombinant DNA molecule. Areas on the slides containing incompletely stretched molecules or elevated levels of background fluorescence are usually skipped when recording images.

The equilibration of high molecular weight DNA in 100 mM NaCl appeared to help in stabilizing the DNA molecules. This effect was more pronounced on larger DNA molecules such as YACs. It should also be pointed out that pipetting the YACs in solution was done only with large opening pipette tips, and reduced to a minimum. Very large DNA molecules such as YACs tend to sediment rapidly. For mixing and redistribution, tubes were flicked very gently with a finger and not vortexed. When using megabase size DNA, mixing was kept to a minimum. The addition of AF-solution to the DNA solution prior to molecular combing is optional. The glycerol might lead to decreased breakage due to reduced Brownian motion, but slows down the drying process. However, the AF-solution should be included when the combing process is observed in the fluorescence microscope. Cumulatively, these measures reduced the breakage of DNA molecules to the point where we found about one full length molecule in every microscopic field of view.

3.2. Mapping of DNA clones along YAC DNA fibers

The choice of colors for a particular target is arbitrary. We typically chose the detection scheme with the highest sensitivity, i.e. digoxigenin-labeled probes detected with rhodamine, for the smallest target. The dual color detection scheme reveals regions of overlap as red/yellow stained regions, while parts of the BAC or YAC DNA fiber such as flanking sequences that do not bind the cosmid, plasmid or the P1 probe appear in green. In the three color scheme, we often detect the counterstained DNA fiber in the blue wavelength interval, the vector specific probes in green and the probes to be mapped in red or green. Again, this is arbitrary and, as one example in this paper shows, the fiber can be detected in green, if so desired.

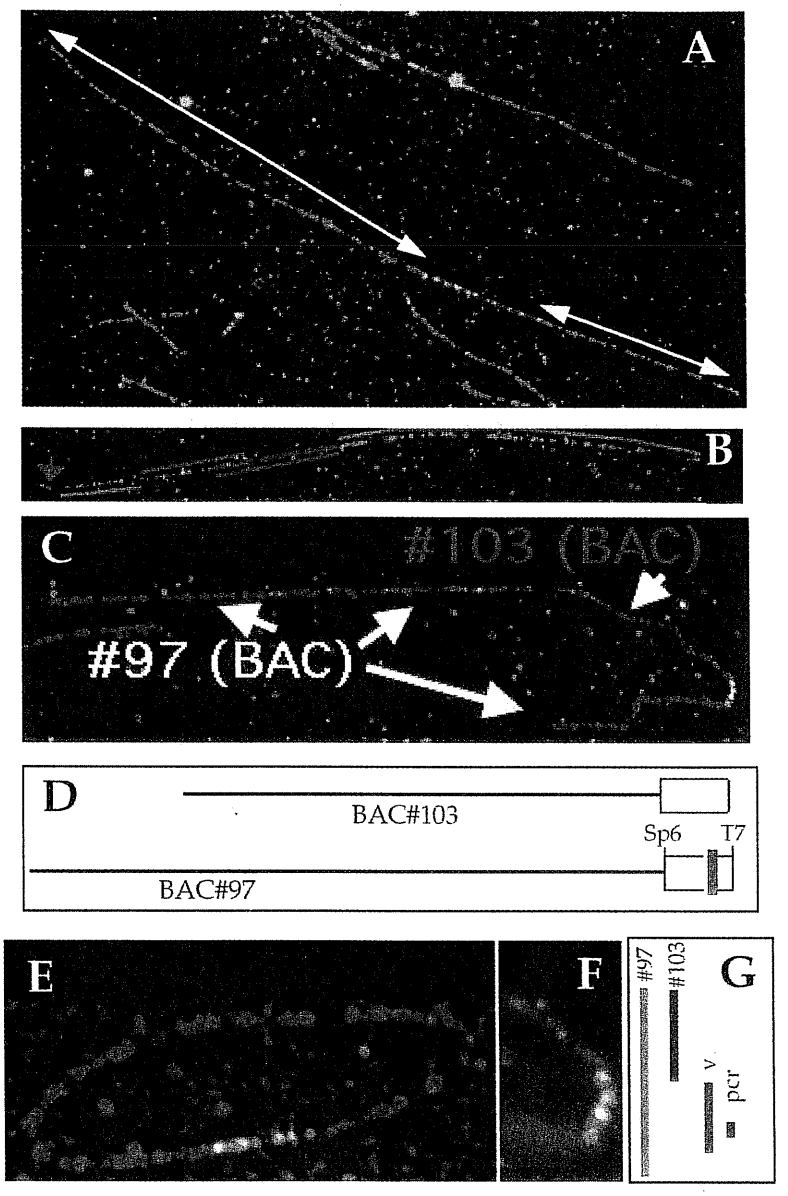

Figure 2. FISH using straightened DNA molecules. The fibers were counterstained by hybridization of probes shown in green (A, B) or blue (C, E, F). The probes to be mapped are displayed in red. (A) Mapping of a P1 clone (shown as red/yellow hybridization domain) along a YAC DNA molecule. The map position is determined by measuring the distances from either end of the YAC molecule as indicated by the arrows. The ends of the YAC molecule (green) are labeled by hybridization of a YAC-vector specific probe (red). (B) Localization of two cosmid clones in a genomic interval defined by a YAC DNA molecule (clone 35B2) and determination of overlap between them. Hybridization domains of the cosmid clones are visualized in red and blue, respectively, while the entire YAC molecules were counterstained by hybridization with a probe that is shown in green. The colored bars indicate the hybridization domains. A probe for the YAC vector (pJS97, red arrow) was included to allow measurement of the distances from the end of the YAC DNA molecule. Numerical values for overlap and distances are given in table I. (C)–(G) Determination of overlap between two linked BAC clones (BAC#97, BAC#103). The images were recorded from either a randomly broken molecule (BAC#97, C) or from a circular DNA molecule (BAC#103, E). Probes specific for the BAC vector are shown in green and a small PCR generated probe for the vector (PCR) is shown in red. The schematic in (D) summarizes the measured overlap and relative orientation of the two clones. An enlarged display showing the hybridization of probes in the vector region of BAC#97 (C) is shown in (F). The probes binding in the region are identified in (G) (#97 = BAC#97, #103 = BAC#103, v = vector, pcr = PCR-made probe for part of the vector).

It should also be noted that Cot-1 DNA, normally added to the hybridization mixture to block repeated DNA sequences during FISH with large-insert DNA probes (Weier *et al* 1995b), is not needed in this protocol, because the repeats are widely distributed along the combed DNA fibers (Weier *et al* 1995b). On the other hand,

we found addition of unlabeled Cot-1 DNA useful to increase signal contrast when mapping relatively small probes such as plasmid clones with inserts of a few kb of genomic DNA (data not shown). Furthermore, the two different denaturation protocols described in the material and methods section lead to identical results.

The localization of single P1 or cosmid probes in intervals defined by collinear YAC molecules is a relatively straightforward application of QDFM. A typical example employing a dual color scheme is shown in figure 2A. We prepared DNA fibers from the ~ 490 kb YAC clone '141G6' and mapped the ~ 80 kb P1 clone '1143'. The YAC molecules were counterstained with a biotinylated YAC probe (detected in green), while the P1 DNA probe was labeled with digoxigenin (dig.) and detected in red. The hybridization mixture also contained dig.-labeled DNA from plasmid pYAC3 to specifically mark the ends of the YAC molecules (small red signals at the ends of the YAC molecules). The relative mapping position along the YAC molecule was determined by measuring the distances of the ends of the red (P1) hybridization domain from the respective YAC ends as indicated in figure 2. In addition, the extent of the P1 hybridization domain as well as the length of the entire molecule was recorded with each DNA fiber. Differences between distance measurements done in triplicate were typically in the order of 1–3 pixels. The P1 '1143' shown in this example was found to map 49.2μm or 114 kb from the proximal end and to extend 35.32μm or 81 kb (Weier *et al* 1995b).

While measurement of all the distances may not be necessary in routine applications of QDFM, it provided valuable information about uniformity of stretching, breaks in the fiber as well as operator errors in the developmental stage of this technique. Relative standard deviations (CVs) were typically in the order of 5–8% of the measured value. Significantly higher CVs provided a simple means to control the analysis procedure. They prompted us to check the data analysis results for operator errors and unsuitable images such as ones taken from broken molecules or insufficiently stretched fibers.

The throughput of QDFM can be increased by using an additional third color and hybridizing two clones simultaneously. The mapping of two cosmid clones onto a YAC molecule using the triple color scheme is shown in figure 2B. The DNAs from cosmid clones 31F3 and 52F2 were labeled with dig. and biotin, respectively, hybridized to fibers made from YAC 35B2 and detected in red and blue, respectively. Probes specific for one end of the YAC molecule (pJS97, red arrow) and FITC-labeled YAC DNA (detected in green) were included in the hybridization mix. The molecule showed overlapping hybridization domains of the cosmid probes, which was expected based on data published by Kawasaki *et al* (1995). Evaluation of only six molecules allowed us to determine the relative mapping positions and overlap between the cosmid clones with an accuracy of a few per cent. The average values and standard deviations listed in table 1 demonstrate the high precision of QDFM. We also determined the accuracy of QDFM by PFGE sizing of clone 52F2. The pulsed field gel showed a single band around 45 kb (data not shown), which agreed well with the size of clone 52F2 as determined by QDFM (43.9 ± 3.7 kb).

Table 1. Determination of overlap between cosmids 31F3 and 52F2.

Length of cosmid 52F2:	43.9 ± 3.7 kb
Length of cosmid 31F3:	40.4 ± 3.1 kb
Overlap between cosmids:	19.2 ± 3.1 kb
Length of vector probe pJS97:	4.6 ± 1.3 kb
Distance between pJS97 and 31F3:	29.0 ± 3.4 kb

3.3. Quantitation of overlap between two collinear DNA molecules

Presently, most physical maps are constructed based on either clone fingerprinting or sequence tagged site (STS) content mapping. An increasing use of QDFM in our laboratory is the determination of overlap extent and direction between two clones that share one or several STSs. The results of one such investigation are illustrated in figure 2C–E. We were interested in determining the overlap between two BAC clones, BAC#97 and BAC#103, which were know to have at least one STS in common. As mentioned above, the BAC vector contains two Not I sites, and we decided to map on either randomly broken or circular molecules. Figure 2C shows the results of mapping BAC#103 (red) on to fibers from clone BAC#97 (blue). The BAC vector part was counterstained by hybridizing the four FITC-labeled and one dig.-labeled PCR products (green and red, respectively). As figure 2C shows, the red hybridization domain of BAC#103 is bordered by the BAC vector (green) on one side. There is no (blue) gap between BAC#103- and vector-specific signals. Blue signals from the insert of BAC#97 are seen flanking the opposite end of the red hybridization domain. The dig.-labeled, red-visualized PCR-derived BAC vector probe, which was known to map closer to the T7 than the Sp6 promoter in the BAC vector, allowed us to determine the orientation of BAC#97. Hybridization of this PCR generated probe is observed on virtually 100% of complete BAC DNA molecules providing us with an additional quality control mechanism for hybridization and detection regardless of clone overlap. The insert of BAC#103 overlaps a total of 74.2 ± 4.2 kb with BAC#97 on the end of the insert near the Sp6 promoter based on the measurement of 10 molecules. The size of the BAC vector in the same 10 images measured 6.6 ± 0.5 kb, which agreed well with the known size of ~ 7 kb. Knowledge of vector orientation allowed it to be indicated in the map shown in figure 2D.

We then hybridized BAC#97 onto fibers prepared from BAC#103 (figure 2E) to determine whether BAC#103 is

fully contained in BAC#97 or extends beyond. Linear as well as circular molecules (figure 2E) indicated that there is no part of BAC#103 that remained unstained by BAC#97 probe or the vector probe, i.e. showed up as blue domain. This proved that BAC#103 is fully contained in BAC#97 as indicated in the map in figure 2D. The circular DNA molecules were found in a wide range of diameters representing different degrees of stretching. The largest circles were stretched to about $\sim 2.3\,\mathrm{kb}\,\mu\mathrm{m}^{-1}$, but smaller, more condensed moecules could also be analysed using the extent of the (green) vector specific 7 kb domain as standard for normalization.

3.4. Localization of small DNA targets

The detection sensitivity of our protocol is sufficient to localize a $\sim$ 1 kb DNA sequence in virtually all molecules that contain it. An example is shown in figure 2C and 2F. We prepared a highly specific probe DNA for a stretch of $\sim$ 1100 bp of BAC vector DNA by *in vitro* DNA amplification using the polymerase chain reaction (PCR) and labeled it with digoxigenin by random priming. As shown for one molecule in figure 2C, dig.-labeled probes bound on virtually all molecules inspected, adding orientational information to the localization of vector sequences. Figure 2F shows an enlarged image of the vector region of the DNA fiber shown in figure 2C. The BAC#97 is counterstained in blue, the BAC#103 probe gave red hybridization signals and most of the vector part is seen as a green domain (figures 2F and G). The $\sim$ 1100 bp dig.-labeled PCR probe gave a clear signal towards one end of the green domain. Similar probes were prepared for the various parts of the P1 vector and applied with equal success.

We have limited experience with detection of DNA sequences smaller than 1 kb. In previously reported experiments, we hybridized cosmid-derived DNA probes to P1 molecules and found that a cross-hybridizing $\sim$ 500 bp sequence common to both cloning vectors gave a very clear signal in about 60% of the molecules (Weier *et al* 1995b).

3.5. Whole chromosome DNA fiber FISH

The interest in using the CEPH/Genethon megaYAC clones, i.e. clones from the library with plate numbers > 730 (Weissenbach *et al* 1992), for mapping and positional cloning is tremendous. This can be attributed to the fact that most of these clones have been placed on integrated genetic and physical maps (Hudson *et al* 1995) and the inserts are relatively large. Many of the megaYAC clones carry an insert larger than 1 Mbp, which complicates DNA fiber mapping for several reasons: the molecules tend to break easier, the number of molecules that can be bound to the slides needs to be low to minimize overlap between DNA fibers and interference during molecular combing, and images of intact molecules have to be recorded in several frames.

As a model for mapping in intervals of 1–2 Mbp, we investigated the application of QDFM for localization of individual lambda clones along intact yeast chromosomes. Several of the chromosomes of the yeast Saccharomyces cerevisiae are between 1 and 2 Mbp in size. An aberrant PFGE-purified yeast (S. cerevisiae) chromosome VII in agarose and DNA from a lambda clone that maps in the centromeric region (ATTC L70464) were provided to us by collaborators (A Admire and T Weinert, University of Arizona). Further information about the lambda clone was obtained from the Stanford University www server at gopher://genome-gopher-.stanford.edu:70/OR72782–72973-/.index/clone.sacchdb/.

According to the database entry, this clone maps on S. cerevisiae chromosome VII at position 481320–497794, suggesting an insert size of 16.5 kb. Sizing of the aberrant chromosome VII by PFGE suggested a size of 1200 kb instead of the normal size of 1088–1120 kb (see the Stanford and ATTC databases). Thus, the aberrant yeast chromosome VII is about 80–100 kb longer that a normal chromosome VII and we were interested in mapping its centromeric DNA sequences.

A typical result of mapping the lambda clone L70464 along a yeast chromosome is shown in figure 3. The DNA fibers were prepared from undigested linear yeast molecules. Following molecular combing, the chromosome (DNA fiber) was counterstained by hybridization of biotinylated total yeast DNA probe detected with avidin-FITC (green). Three overlapping images were recorded using a 40× microscope objective. The hybridization domain of the dig.-labeled lambda probe (red, arrow in figure 3A) was readily visible. This mapping experiment was particularily difficult due to the lack of probes marking the ends of the molecule, and we decided to choose the largest molecules on the slide for analysis. Measurements of the distances of the lambda hybridization domain from the ends of the fiber shown in figure 3A suggested that the molecule had a total length of about 1180 kb. Our calculations were based on the known pixel spacing of $0.17\,\mu\mathrm{m}$/pixel for the 40× lens and a linear factor of $2.3\,\mathrm{kb}\,\mu\mathrm{m}^{-1}$. The lambda clone hybridization domain thus extended for 18.0 kb and the distances from the two ends were 523.6 kb and 638.8 kb, respectively. The exact position of the S. cerevisiae centromere within the lambda clone was not known to us, but the results were in excellent agreement with the map of the normal (1088–1120 kb) S. cerevisiae chromosome VII with left and right arm lengths of 450–500 kb and 600–631 kb, respectively. Furthermore, the results indicated that the left arm of this aberrant chromosome VII must harbor the extra 60–80 kb. Figure 3B shows a schematic representation of the lambda probe and its approximate mapping position.

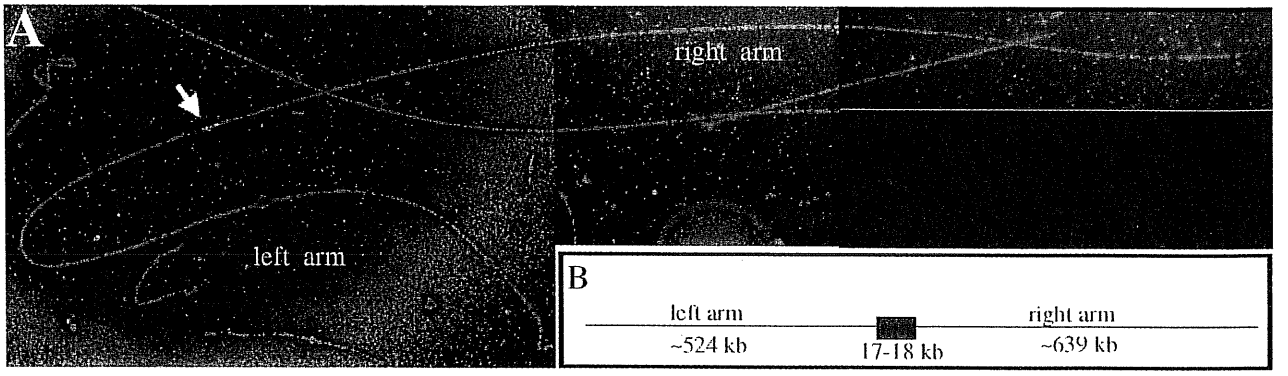

Figure 3. Whole chromosome DNA fiber mapping. (A) Yeast chromosomes VII of 1200 kb, i.e. slightly larger than a normal S. cerevisiae chromosome VII (1120 kb), were isolated by PFGE and stretched on APS-slides. We then mapped a lambda clone of ~ 17 kb containing the centromeric region (red domain, arrow). The yeast chromosomes were counterstained by hybridization with a total yeast DNA probe (green). The image had to be recorded in three frames, which were later cropped and joined using Adobe Photoshop. (B) Schematic diagram indicating the map position and sizes of the lambda clone and left and right chromosome arms.

3.6. Discussion

Possibly the two most frequently asked questions in physical map assembly are related to the distance of two markers or clones and the extent of overlap between clones. QDFM is in a favorable position to answer both these questions. When a high molecular weight DNA molecule such as a YAC is available that spans the region of interest, the position of each clone within this genomic interval can be determined easily as shown in figures 2A and B. The distance between two clones becomes readily apparent by hybridizing probes from both clones onto the DNA fiber. This way, gaps between contigs can be quantitated accurately and rapidly (Weier *et al* 1995b). Since it is known that some YAC clones have a high propensity to delete parts of their insert, the data have to be looked at with caution. Although this is not a guarantee that our YACs have no deletions, we typically size and purify all YAC DNA molecules and compare the sizes determined by PFGE with sizes listed for the same clone in publicly accessible databases and the literature. We typically use the gel slice containing the largest YAC molecule for our studies. If YAC clones from our copies of the library are significantly smaller that expected, we try to obtain new clones. If mapping data obtained by QDFM using YACs are inconsistent with other data (for example, clones cannot be localized because the YAC has a deletion), the overlap can be determined by mapping one clone onto the other as shown for the BAC clone in figures 2C and E.

In general, the instability seems to increase with the size of the YAC insert, and becomes a more severe issue when using clones from the CEPH megaYAC library. The usefulness of QDFM extends well into the megabasepair range. We have performed several mapping studies using yeast and YAC DNA molecules larger than 1 megabasepairs (Mbp). Working with such large molecules (a 1 Mbp molecule is ~ 400 μm long) gives rise to additional complications, most notably that the molecules break more easily and images from straight molecules cannot be recorded in a single camera frame. Nevertheless, QDFM does not require many molecules for accurate measurements. We deposit several thousand molecules per slide, so that we typically have no problems finding the desired number of about ten full length molecules for measurements, and QDFM has already proven to be most useful to map relatively small lambda and P1/PAC clones in such large genomic intervals.

The mechanisms of binding purified DNA molecules to amino-silanated glass or mica are not yet understood (Weier *et al* 1995b, Hu *et al* 1996), but studies using linear double-stranded DNA molecules such as phage DNA or yeast artificial chromosomes showed preferential binding of the ends of the DNA molecules to the solid substrate (Bensimon *et al* 1994, Weier *et al* 1995b). Silanation of the glass slides greatly affects the speed and density of binding of DNA molecules to the surface. Repeated experiments in our laboratory have shown that some DNA molecules bind on one end during assembly of the slide–coverslip sandwich and are stretched by the hydrodynamic force pointing outwards, while most molecules bind later and are stretched during the drying process. Initially, we assumed a very low binding efficiency for circular DNA molecules, and a somewhat higher binding efficiency when the DNA circles were nicked on one strand. We found that our alkaline lysis procedure generates single-stranded nicks frequently in DNA molecules, as suggested by the deposition of numerous such circular DNA molecules (figure 1C).

The use of circular DNA requires normalization of DNA molecules, which appear not to stretch as well as linear molecules. We therefore investigated the use of randomly broken linear DNA molecules for QDFM. These molecules seem to be the preferred mapping templates

for a wide variety of cloning vectors, because their use circumvents the above mentioned problems with enzymatic linearization. The inclusion of just few vector-specific probes in the hybridization probe mixture allows easy interpretation of images and mapping data obtained from randomly broken molecules.

The sensitivity of QDFM using standard FISH protocols without modification is sufficient to localize clone overlap of at least one 1 kb. Typically, we immuno-cytochemically amplify the signals to the point where they can be detected by eye and recorded on film, if so desired. Improvement in the FISH and detection protocols may increase the reliability of detecting sub-kb targets. Although not shown in this paper, we also have experience with molecular combing and fiber FISH on short DNA molecules such as plasmids of 7 kb. In our laboratory, fibers made from such small clones are typically used to test probes after labeling.

Our results demonstrate that QDFM can localize ~ 17–18 kb lambda clones on megabase size S. cerevisiae chromosomes with a precision of a few kb. This can provide valuable information for the process of yeast chromosome sequencing (Galibert *et al* 1996), because mapping a few clones or sequencing templates along a yeast molecule, for example, will provide valuable checkpoints for sequence assembly.

4. Outlook

Quantitative DNA fiber mapping will find widespread application in large scale mapping and production DNA sequencing, once it has been demonstrated that APS-slide preparation, molecular combing, fiber detection and analysis scale up to the need of such environments. The development of high-throughput systems and algorithms for automated analysis of QDFM slides will therefore have high priority. Modifications to the protocols like the use of infrared fluorescent dyes such as CY5 for detection of additional probes and full utilization of the high sensitivity of electronic detectors (cooled CCD or ISIT cameras) may greatly increase the throughput of QDFM mapping. Furthermore, simplified DNA purification and molecular combing protocols will greatly accelerate the processes of general acceptance and utilization of this technique.

The QDFM technique provides valuable tools for quality control in larger mapping or sequencing projects. The technology is applicable for detection of problematic clones, such as chimeric sequencing templates or clones containing duplicated regions of the genome (data not shown). Finally, knowledge about the localization and orientation of contigs can be obtained easily allowing more parallelism in large scale DNA sequencing projects.

We have not systematically investigated the limits of detection, but it appears that the technology is ready for routine mapping of probes of 1 kb or larger. Signal amplification systems combined with the image analysis capability to align and average many molecules might soon allow researchers to localize targets as small as 50–100 bp and, for the first time, visualize sequence tagged sites (STSs) or exons directly on single DNA molecules.

Acknowledgments

Supported by a grant from the Director, Office of Energy Research, Office of Health and Environmental Research, US Department of Energy, under contract DE-AC-03-76SF00098 and a fellowship from the Deutsche Forschungsgemeinschaft (to TD).

References

Bensimon A, Simon A, Chiffaudel A, Croquette V, Heslot F and Bensimon D 1994 Alignment and sensitive detection of DNA by a moving interface *Science* **265** 2096–8

Birnboim H C and Doly J 1979 A rapid alkaline extraction procedure for screening recombinant plasmid DNA *Nucl. Acid Res.* **7** 1513–23

Florijn R J, Bonden L A J, Vrolijk H, Wiegant J, Vaandrager J-W, Baas F, den Dunnen J T, Tanke H J, van Ommen G-J B and Raap AK 1995 High-resolution DNA Fiber-FISH for genomic DNA mapping and colour bar-coding of large genes *Human Mol. Genet.* **4** 831–6

Galibert F *et al* 1996 Complete nucleotide sequence of Saccharomyces cerevisiae chromosome X *EMBO J.* **15** 2031–49

Glazer A N and Rye H S 1992 Stable dye-DNA intercalation complexes as reagents for high-sensitivity fluorescence detection *Nature* **359** 859–61

Haaf T and Ward DC 1994 Structural analysis of α-satellite DNA and centromere proteins using extended chromatin and chromosomes *Human Mol. Genet.* **3** 629–33

Heiskanen M, Karhu R, Hellsten E, Peltonen L, Kallioniemi O P and Palotie A 1994 High resolution mapping using fluorescence *in situ* hybridization to extended DNA fibers prepared from agarose-embedded cells *BioTechniques* **17** 928–33

Heng H H Q, Squire J and Tsui L C 1992 High-resolution mapping of mammalian genes by *in situ* hybridization to free chromatin *Proc. Natl Acad. Sci. USA* **89** 9509–13

Hu J, Wang M, Weier H-U G, Frantz P, Kolbe W, Olgletree D F and Salmeron M 1996 Imaging of single extended DNA molecules on flat (aminopropyl)–triethoxysilane-mica by atomic force microscopy *Langmuir* **12** 1697–700

Hudson T J *et al* 1995 An STS-based map of the human genome *Science* **270** 1945–54

Ioannou P A, Amemiya C T, Garnes J, Kroisel P M, Shizuya H, Chen C, Batzer M and de Jong P J 1994 A new bacteriophage P1-derived vector for the propagation of large human DNA fragments *Nature Genet.* **6** 84–9

Johnson G D and Nogueira Araujo G M 1981 A simple method of reducing the fading of immunofluorescence during microscopy *J. Immunol. Meth.* **43** 349–51

Kawasaki K, Minoshima S, Schooler K, Kudoh J, Asakawa S, de Jong P J and Shimizu N 1995 The organization of the human immunglobulin λ locus *Genome Res.* **5** 125–35

Lawrence J B, Carter K C and Gerdes M J 1992 Extending the capabilities of interphase chromatin mapping *Nature Genet.* **2** 171–2

Mascio L N, Verbeek P W, Sudar D, Kuo W L and Gray J W 1995 Semiautomated DNA probe mapping using digital imaging microscopy: I. System development *Cytometry* **19** 51–69

Parra I and Windle B 1993 High resolution visual mapping of stretched DNA by fluorescent hybridization *Nature Genet.* **5** 17–21

Patil N, Peterson A, Rothman A, de Jong P J, Myers R M and Cox D R 1994 A high resolution physical map of 2.5 Mbp of the Down syndrome region on chromosome 21 *Human Mol. Genet.* **3** 1811–7

Rosenberg C, Florijn R J, Van de Rijke F M, Blonden L A, Raap T K, Van Ommen G J and Den Dunnen J T 1995 High resolution DNA fiber-fish on yeast artificial chromosomes: direct visualization of DNA replication *Nature Genet.* **10** 477–9

Shepherd N S, Pfrogner B D, Coulby J N, Ackerman S L, Vaidyanathan G, Sauer R H, Balkenhol T C and Sternberg N 1994 Preparation and screening of an arrayed human genomic library generated with the P1 cloning system *Proc. Natl Acad. Sci. USA* **91** 2629–33

Sherman F, Fink G R and Hicks J B 1986 *Laboratory Course Manual for Methods in Yeast Genetics* (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory)

Shizuya H, Birren B, Kim U J, Mancino V, Slepak T, Tachiiri Y and Simon M 1992 Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector *Proc. Natl Acad. Sci. USA* **89** 8794–7

Suto Y, Tokunaga K, Watanabe Y and Hirai M 1996 Visual demonstration of the organization of the human complement C4 and 21-hydroxylase genes by high-resolution fluorescence *in situ* hybridization *Genomics* **33** 321–4

Tocharoentanaphol C, Cremer M, Schrock E, Blonden L, Kilian K, Cremer T and Ried T 1994 Multicolor fluorescence *in situ* hybridization on metaphase chromosomes and interphase halo-preparations using cosmid and YAC clones for the simultaneous high resolution mapping of deletions in the dystrophin gene *Human Genet.* **93** 229–35

Weier H-U G, Rhein A P, Shadravan F, Collins C and Polikoff D 1995a Rapid physical mapping of the human *trk* proto-oncogene (NTRK1) gene to human chromosome 1q21-22 by P1 clone selection, fluorescence *in situ* hybridization (FISH) and computer-assisted microscopy *Genomics* **26** 390–3

Weier H-U G, Wang M, Mullikin J C, Zhu Y, Cheng J-F, Greulich K M, Bensimon A and Gray J W 1995b Quantitative DNA fiber mapping *Human Mol. Genet.* **4** 1903–10

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G and Lathrop M 1992 A second-generation linkage map of the human genome *Nature* **359** 794–801

Wiegant J, Kalle W, Mullenders L, Brookes S, Hoovers J M, Dauwerse J G, van Ommen G J and Raap A K 1992 High-resolution *in situ* hybridization using DNA halo preparations *Human Mol. Genet.* **1** 587–91